

Article

Unraveling the Mitogenomic Characteristics and Phylogenetic Implications of *Leuciscus merzbacheri* (Zugmayer, 1912), an Endangered Fish in the Junggar Basin of Xinjiang, Northwest China

Yan Sun and Tianyan Yang * 

Fishery College, Zhejiang Ocean University, Zhoushan 316022, China; sy1249032764@163.com

* Correspondence: hellojelly1130@163.com

Abstract: Background: *Leuciscus merzbacheri* is a rare and endangered fish in Xinjiang, China. As a representative species of the fauna in the Junggar Basin, it is of high economic and scientific value. The genetic data are still limited, and the mitochondrial genomic characteristics remain unexplored. Methods: A high-throughput sequencing method was used to obtain the complete mitogenome of *L. merzbacheri*. Results: The full length of the circular DNA was 16,609 bp, and it consisted of 13 protein-coding genes (PCGs), 22 tRNAs, 2 rRNAs and 2 non-coding regions. The overall nucleotide compositions of both the mitogenome and PCGs showed an obvious AT preference with percentages of 54.20% and 53.60%, respectively. Three commonly used amino acids were Leu (16.43%), Ala (8.95%) and Thr (7.85%) in turn. All tRNAs could form the typical clover structures excluding tRNA-Ser^{AGY}. The presumed secondary structures of two rRNAs contained several stem-loop domains, and the structure of 12S rRNA seemed to be more stable than that of 16S rRNA. Extended termination sequence regions (ETASs), central conserved regions (CSB-F, CSB-E and CSB-D), and conserved sequence regions (CSB-1, CSB-2 and CSB-3) were identified in the control region. The phylogenetic tree showed that *L. merzbacheri* was recovered with strong supports as a sister to the other members of the genus. The location in the outermost branch implied that it might be a relatively ancient species among its congeners. Conclusions: This study would complement the genetic data on *L. merzbacheri* and contribute to a better understanding of molecular evolution in *Leuciscus* as well.

Keywords: Illumina sequencing; mitogenome; phylogenic relationship; Leuciscinae



Citation: Sun, Y.; Yang, T. Unraveling the Mitogenomic Characteristics and Phylogenetic Implications of *Leuciscus merzbacheri* (Zugmayer, 1912), an Endangered Fish in the Junggar Basin of Xinjiang, Northwest China. *Genes* **2024**, *15*, 1284. <https://doi.org/10.3390/genes15101284>

Academic Editors: Jiayong Zhang and Jiasheng Hao

Received: 12 September 2024

Revised: 24 September 2024

Accepted: 24 September 2024

Published: 30 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Leuciscus fishes are widely distributed in rivers and lakes in cold-temperate plain areas of Eurasia [1]. The number of species or subspecies within the genus *Leuciscus* distributed in China nearly accounts for 25% of the world's daces [2]. *Leuciscus merzbacheri* is an indigenous fish unique to the Xinjiang Uygur Autonomous Region and only inhabits in the Junggar Basin in the north of Tianshan Mountains, which makes it a landmark species of the fish fauna in this basin [3]. It formerly was the major economical fish of the local areas [3]. However, because of the ecological environment changes and the strengthening of human activities, its habitat is seriously destroyed and the population resource is shrinking continuously [3,4]. Thus, early in 1998, this native fish was listed as a vulnerable species in the “China Red Data Book of Endangered Animals (Pisces)” [5]. Nowadays, it has been considered first-class protected aquatic wildlife in Xinjiang and is also included in “The IUCN Red List of Threatened Species” as a vulnerable species [6]. So far, studies on *L. merzbacheri* have mostly concentrated on the fields of morphological and biological characteristics [7–10], resource distribution [3,11] and artificial reproduction and protection [12,13]. A finite number of genetic references is available for this rare and

endangered fish species, and only the chromosome karyotype and tissue-specific expression of isozymes have been reported until now [14,15]. At the molecular level, only a draft mitogenome and some mitochondrial and nuclear gene fragments, including mitochondrial cytochrome b (Cyt *b*), the control region and the β -actin gene promoter, were sequenced and analyzed [16–19], which seriously prevented the genetic conservation and utilization of this critically valuable species.

The mitochondrion is an important place for cell respiration, and the adenosine triphosphate (ATP) produced by it can supply energy for life activities [20]. This semiautonomous organelle possesses the only genetic material outside the nucleus and has the capabilities of independent replication, transcription and translation [21,22]. Mitochondrial DNA (mtDNA) is the physical embodiment of the genetic information encoded in the mitochondrion, and this covalently closed double-stranded circular molecule owns specificities, such as maternal inheritance, a lack of recombination, a high copy number and a rapid evolutionary rate [23]. It has become an effective tool for population genetics, molecular phylogenetics, species identification and forensic science [24–26]. However, the limited mitochondrial DNA data confine the genetic and taxonomic studies on *L. merzbacheri*. Only Hai et al. estimated the divergence time of Leuciscinae in Xinjiang based on mtDNA Cyt *b* sequences, which revealed that *Leuciscus* originated at the Miocene, and the taxonomic relationship of *L. merzbacheri* was relatively distant from *L. baicalensis*, *L. idus* and *L. waleckii* [27]. Furthermore, its taxonomic status within the genus *Leuciscus* is also confused because of higher species diversity and morphological similarity and even introgressive hybridization in natural homoploid fishes [28–30].

In this study, the complete mtDNA genome of *L. merzbacheri* was sequenced, assembled and annotated. Meanwhile, the phylogenetic tree was constructed using the complete mitogenome sequences of the Leuciscinae species published in the GenBank database to provide reference for its classification status. The results will help to promote future studies on the conservation genetics and evolutionary biology of this native and threatened species and other fishes in *Leuciscus*.

2. Materials and Methods

2.1. Sampling, DNA Extraction and High-Throughput Sequencing

The sample of *L. merzbacheri* was collected in October 2021 from a tributary of the Manas River, which originates from the northern slope of the Tianshan Mountains in Xinjiang, China (Figure 1). A small amount of pectoral fin tissue was clipped and stored in absolute ethanol immediately. The genomic DNA was extracted using the standard phenol–chloroform method [31]. The integrity and purity of DNA were detected via 1% agarose gel electrophoresis and with a Qubit fluorometer, respectively. The qualified template DNA was sheared into fragments of a specific size (300–500 bp) via ultrasonic mechanical interruption (Covaris M220, Woburn, MA, USA), and then, the fragmented DNA was processed through purification, terminal repair, A-tail and sequencing adaptor addition and fragment size selection using agarose gel electrophoresis. Subsequently, the appropriate DNA fragments were amplified via PCR to construct the paired-end library. The prepared library was sequenced on an Illumina HiSeq based on sequencing by synthesis (SBS) technology.

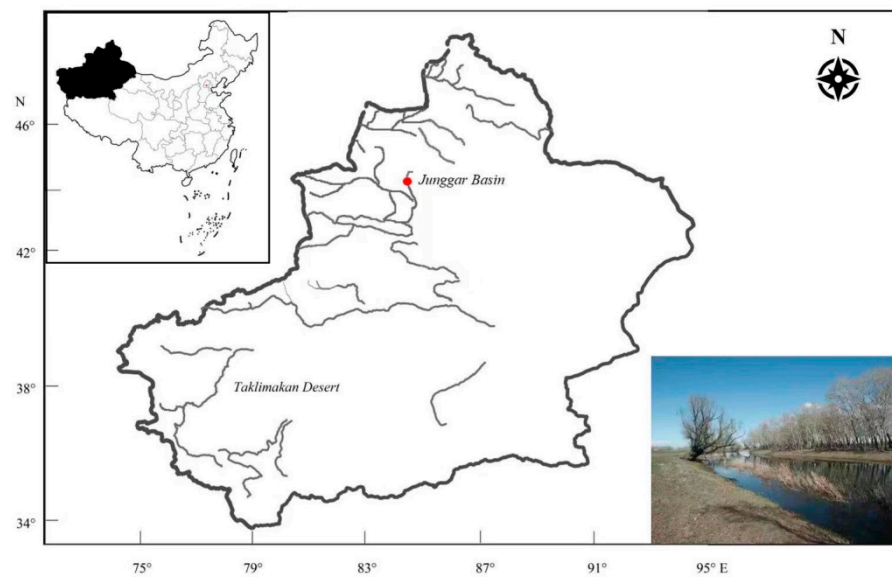


Figure 1. Sampling site of *L. merzbacheri* in this study.

2.2. Mitochondrial Genome Assembly and Annotation

Prior to assembly, the raw data were purified by removing adaptors and primers, as well as low quality reads with a quality score lower than 20 ($Q < 20$) by using Cutadapt 4.4 software [32]. MITObim 1.9.1 software was applied to assemble the mitochondrial genome with default parameter settings based on the clean reads mapped to the reference mitogenome of *L. baicalensis* (GenBank accession No. KF673863) [33]. The obtained mtDNA sequence was annotated using the web-based tool MitoAnnotator (<http://mitofish.aori.u-tokyo.ac.jp/annotation/input.html>; accessed on 20 February 2024). The mitogenome map of *L. merzbacheri* was plotted using OGDRAW software 1.3.1 [34]. The entire mitogenome sequence was deposited into the Nucleotide database of the National Center for Biotechnology Information (NCBI) under the accession number OR992052.

2.3. Mitogenome Sequence Analysis

The online tools tRNAscan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE/>; accessed on 5 March 2024) and tRNADB (<http://rna.tbi.univie.ac.at/forna/>; accessed on 6 March 2024) were utilized to locate and visualize the cloverleaf structures of tRNA using the default search modes, respectively. The secondary structures of two rRNAs were prognosticated using the R2DT tool of the RNACentral website (<https://rnacentral.org/> (accessed on 10 March 2024)). The location of origin of the L-strand replication region (O_L) was identified via sequence homology alignment, and its secondary structure was predicted online with Mfold (<http://www.unafold.org/>; accessed on 15 March 2024). Nucleotide sequences were aligned with MUSCLE 3.8 [35]. MEGA 6.0 software was used to analyze the nucleotide content, amino acid composition and relative synonymous codon usage (RSCU) of the protein-coding gene (PCG) sequences [36]. The formulas $AT\text{-skew} = (A - T)/(A + T)$ and $GC\text{-skew} = (G - C)/(G + C)$ were used to describe the overall patterns of nucleotide composition in mtDNA sequences [37].

2.4. Phylogenetic Analysis

A total of thirty-three complete mtDNA sequences of Leuciscinae species were downloaded from the GenBank database for phylogenetic inference (Table S1). In addition, largemouth bass *Micropterus salmoides* (GenBank accession No. OL339398) belonging to the order Perciformes and family Cehtrachidae was adopted as the outgroup taxa of the evolutionary tree. The trimmed PCG sequences (without stop codons) were concatenated into a matrix with FASconCAT-G 1.04 software [38]. The same optimal nucleotide substitute model “GTR + I + G” for the maximum likelihood (ML) and Bayesian inference (BI)

analyses was selected by Modeltest 3.7 software under the Akaike information criterion (AIC) and Bayesian information criterion (BIC), respectively [39]. The ML phylogenetic tree was constructed using PhyML 3.0 software, and the tree topology was evaluated with 1000 bootstrap replicates [40]. The BI phylogenetic tree was constructed using MrBayes 3.2.6 software, with one cold and three heated Markov Chain Monte Carlo (MCMC) chains running for 1,000,000 generations sampled every 1000 generations, and the first 25% of the samples were discarded as burn in [41].

3. Results

3.1. Mitogenome Structure and Nucleotide Composition

The complete mitogenome of *L. merzbacheri* was 16,609 bp in length, including 13 PCGs, 22 transfer RNAs (tRNAs), 2 ribosomal RNAs (12S rRNA and 16S rRNA) and 2 non-coding regions (D-loop and O_L) (Table 1, Supplementary Figure S1). Among the 37 genes, only 9 genes (tRNA-Pro, tRNA-Glu, *ND6*, tRNA-Ser^{UCN}, tRNA-Tyr, tRNA-Cys, tRNA-Asn, tRNA-Ala and tRNA-Gln) were located on the light strand (L-strand), and the remainders were encoded by the heavy strand (H-strand). In particular, all PCGs except for *ND6* were situated on H-strand. A total of 7 gene overlaps (2–7 bp) and 11 gene spacers (1–13 bp) were detected in the whole mitochondrial genome. The largest overlaps were between *ATP8* and *ATP6*, as well as *ND4L* and *ND4*. Meanwhile, the longest intergenic gaps appeared between tRNA-Asp and *COII* with the length of 13 bp.

Table 1. Structural characteristics of the mitochondrial genome of *L. merzbacheri*.

Genes	Strand	Initiation Site	Termination Site	Length	Spacer (+)/Overlap (–)	Initiation Codon	Termination Codon	Anti-Codon
tRNA-Phe	H	1	69	69	0			GAA
12S rRNA	H	70	1028	959	0			
tRNA-Val	H	1029	1100	72	0			UAC
16S rRNA	H	1101	2791	1691	0			
tRNA-Leu ^{UUR}	H	2792	2867	76	0			UAA
<i>ND1</i>	H	2868	3842	975	–2	ATG	TAG	
tRNA-Ile	H	3847	3918	72	4			GAU
tRNA-Gln	L	3917	3987	71	–2			UUG
tRNA-Met	H	3989	4057	69	1			CAU
<i>ND2</i>	H	4058	5102	1045	0	ATG	T	
tRNA-Trp	H	5103	5173	71	0			UCA
tRNA-Ala	L	5175	5243	69	1			UGC
tRNA-Asn	L	5245	5317	73	1			GUU
O _L	H	5318	5354	37	0			
tRNA-Cys	L	5352	5418	67	–3			GCA
tRNA-Tyr	L	5420	5490	71	1			GUA
<i>COI</i>	H	5492	7042	1551	1	GTG	TAA	
tRNA-Ser ^{UCN}	L	7043	7113	71	0			UGA
tRNA-Asp	H	7117	7190	74	3			GUC
<i>COII</i>	H	7204	7894	691	13	ATG	T	
tRNA-Lys	H	7895	7970	76	0			UUU
<i>ATP8</i>	H	7972	8136	165	1	ATG	TAG	
<i>ATP6</i>	H	8130	8812	683	–7	ATG	TA	
<i>COIII</i>	H	8813	9596	784	0	ATG	T	
tRNA-Gly	H	9597	9668	72	0			UCC
<i>ND3</i>	H	9669	10,017	349	0	ATG	T	
tRNA-Arg	H	10,018	10,087	70	0			UCG
<i>ND4L</i>	H	10,088	10,384	297	0	ATG	TAA	
<i>ND4</i>	H	10,378	11,759	1382	–7	ATG	TA	

Table 1. Cont.

Genes	Strand	Initiation Site	Termination Site	Length	Spacer (+)/Overlap (-)	Initiation Codon	Termination Codon	Anti-Codon
tRNA-His	H	11,760	11,828	69	0			GUG
tRNA-Ser ^{AGY}	H	11,829	11,897	69	0			GCU
tRNA-Leu ^{CUN}	H	11,899	11,971	73	1			UAG
ND5	H	11,972	13,807	1836	0	ATG	TAG	
ND6	L	13,804	14,325	522	-4	ATG	TAA	
tRNA-Glu	L	14,326	14,394	69	0			UUC
Cyt <i>b</i>	H	14,399	15,539	1141	4	ATG	T	
tRNA-Thr	H	15,540	15,611	72	0			UGU
tRNA-Pro	L	15,611	15,680	70	-1			UGG
D-loop	H	15,681	16,609	929	0			

H: heavy-strand; L: light-strand; the negative number indicates the overlap, and superscript means the codon family.

As a remarkable feature of animal mitochondrial genomes, the asymmetry in nucleotide composition of the whole mtDNA was discovered as follows: A 27.90%, G 18.68%, T 26.30% and C 27.12%, showing an obvious A+T bias (54.20%). The patterns of nucleotide distribution throughout the whole mitogenome are exhibited in Supplementary Table S2. The results showed a positive AT-skew value (0.0295) and a negative GC-skew value (-0.1843), respectively. The GC-skew values of 13 PCGs and 2 rRNAs were all less than zero, which illustrated that the nucleotide compositions were heavily anti-bias toward guanine. A conspicuous anti-T bias was observed in *ND6*, which was quite unlike other PCGs and might be related to different coding strands. The D-loop region had the most obvious anti-GC skew of all, accounting for 62.76% of A+T content, and the AT proportion in *ATP8* was also very high, reaching 61.21%. However, only the A+T contents of *ND2* and *ND6* were lower than those of G+C contents, showing a greater abundance of cytosine.

3.2. Protein-Coding Genes, Amino Acids and Codon Usage Pattern

The mitochondrial genome of *L. merzbacheri* contained 13 PCGs with a total length of 11,421 bp, which occupied 68.76% of the entire mitogenome. Twelve PCGs were located on the H-strand, but the *ND6* gene was the only exception, which was encoded on the opposite strand. There were two types of initiator codons used by PCGs, in which ATG most frequently appeared, and only *COI* started with GTG. Five genes (*COI*, *ND1*, *ND4L*, *ND6* and *ATP8*) employed the conventional stop codons TAA and TAG. *ATP6* and *ND4* used TA- as the terminal codon, while five genes (*ND2*, *COII*, *COIII*, *ND3* and *Cytb*) used T-- as the terminal codon.

MEGA6.0 software was used to obtain the RSCU values of 13 PCGs and analyze the codon usage preferences (Table 2). The codon CGA(R) had the highest frequency of RSCU (1.89) with the count number of 87, followed by the codon CCC(P), while the lowest RSCU value 0.32 was found in UCG(S). In general, the numbers of NNT (16) and NNC (16) codons were higher than NNA (14) and NNG (14) codons in the mtDNA PCGs of *L. merzbacheri*, which was consistent with the nucleotide compositions of all 13 PCGs. The amino acid composition was analyzed, and the higher coding frequencies of PCGs were Leu (16.43%), Ala (8.95%) and Thr (7.85%), respectively (Supplementary Figure S2). Among the six available codons for the amino acid leucine, the CUA codon had the highest RSCU value (1.84) within the coding regions analyzed in *L. merzbacheri*.

Table 2. Frequency of codon usage in 13 PCGs of the *L. merzbacheri* mitogenome.

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	95	0.84	UCU(S)	48	1.20	UAU(Y)	55	0.97	UGU(C)	14	1.08
UUC(F)	132	1.16	UCC(S)	59	1.47	UAC(Y)	58	1.03	UGC(C)	12	0.92
UUA(L)	103	0.99	UCA(S)	64	1.59	UAA(*)	0	0.00	UGA(W)	85	1.43
UUG(L)	45	0.43	UCG(S)	13	0.32	UAG(*)	0	0.00	UGG(W)	34	0.57
CUU(L)	118	1.13	CCU(P)	33	0.62	CAU(H)	27	0.52	CGU(R)	7	0.37
CUC(L)	107	1.03	CCC(P)	77	1.44	CAC(H)	76	1.48	CGC(R)	15	0.79
CUA(L)	191	1.84	CCA(P)	82	1.53	CAA(Q)	68	1.39	CGA(R)	36	1.89
CUG(L)	60	0.58	CCG(P)	22	0.41	CAG(Q)	30	0.61	CGG(R)	18	0.95
AUU(I)	161	1.22	ACU(T)	52	0.70	AAU(N)	45	0.79	AGU(S)	17	0.42
AUC(I)	102	0.78	ACC(T)	106	1.42	AAC(N)	69	1.21	AGC(S)	40	1.00
AUA(M)	88	1.02	ACA(T)	105	1.41	AAA(K)	48	1.23	AGA(*)	0	0.00
AUG(M)	85	0.98	ACG(T)	35	0.47	AAG(K)	30	0.77	AGG(*)	0	0.00
GUU(V)	59	0.90	GCU(A)	56	0.66	GAU(D)	32	0.82	GGU(G)	34	0.55
GUC(V)	63	0.96	GCC(A)	149	1.75	GAC(D)	46	1.18	GGC(G)	66	1.06
GUA(V)	98	1.49	GCA(A)	103	1.21	GAA(E)	63	1.24	GGA(G)	84	1.35
GUG(V)	43	0.65	GCG(A)	32	0.38	GAG(E)	39	0.76	GGG(G)	64	1.03

Preference codons are indicated in bold. Asterisk means the stop codon in mitochondrial DNA.

3.3. Characterization and Structural Prediction of tRNAs and rRNAs

Twenty-two tRNAs were interspersed among the mitochondrial genes of *L. merzbacheri*, with the length ranging from 67 to 76 bp. Most tRNAs were encoded by the H-strand, except for tRNA-Cys, tRNA-Ala, tRNA-Glu, tRNA-Pro, tRNA-Gln, tRNA-Tyr, tRNA-Ser^{UCN} and tRNA-Asn. There were a total of 8 gene overlaps or spacers among the 22 tRNA genes. The largest intergenic spacer and overlap were between tRNA-Ser^{UCN} and tRNA-Asp, as well as between tRNA-Ile and tRNA-Gln, respectively. When predicting the secondary structures of tRNAs, it was found that only tRNA-Ser^{AGN} did not form the typical clover structure due to the absence of a dihydrouracil (DHU) arm, while the remaining tRNA genes were all clover shaped (Figure 2). In addition, a total of 46 noncanonical base pairs were detected in 20 tRNAs excluding tRNA-Leu^{CUN} and tRNA-Lys, of which the most abundant type of single-base mismatch was G-U.

The two rRNA genes, 12S rRNA (959 bp) and 16S rRNA (1691 bp), were typically separated by tRNA-Val, and both of them were encoded by the H-strand. The A+T content of 16S rRNA (54.64%) was slightly greater than that of 12S rRNA (50.99%). The secondary structure prediction and visualization of two rRNAs are displayed in Supplementary Figure S3. It was found that 12S rRNA consisted four domains, of which domain I and domain II were variable regions, whereas domain III and domain IV were conserved regions [42]. However, the secondary structure of 16S rRNA contained six domains, and each domain was distinguished by a single-strand sequence. The first four domains (I–IV) were variable regions, while the last two domains (V–VI) were conserved regions.

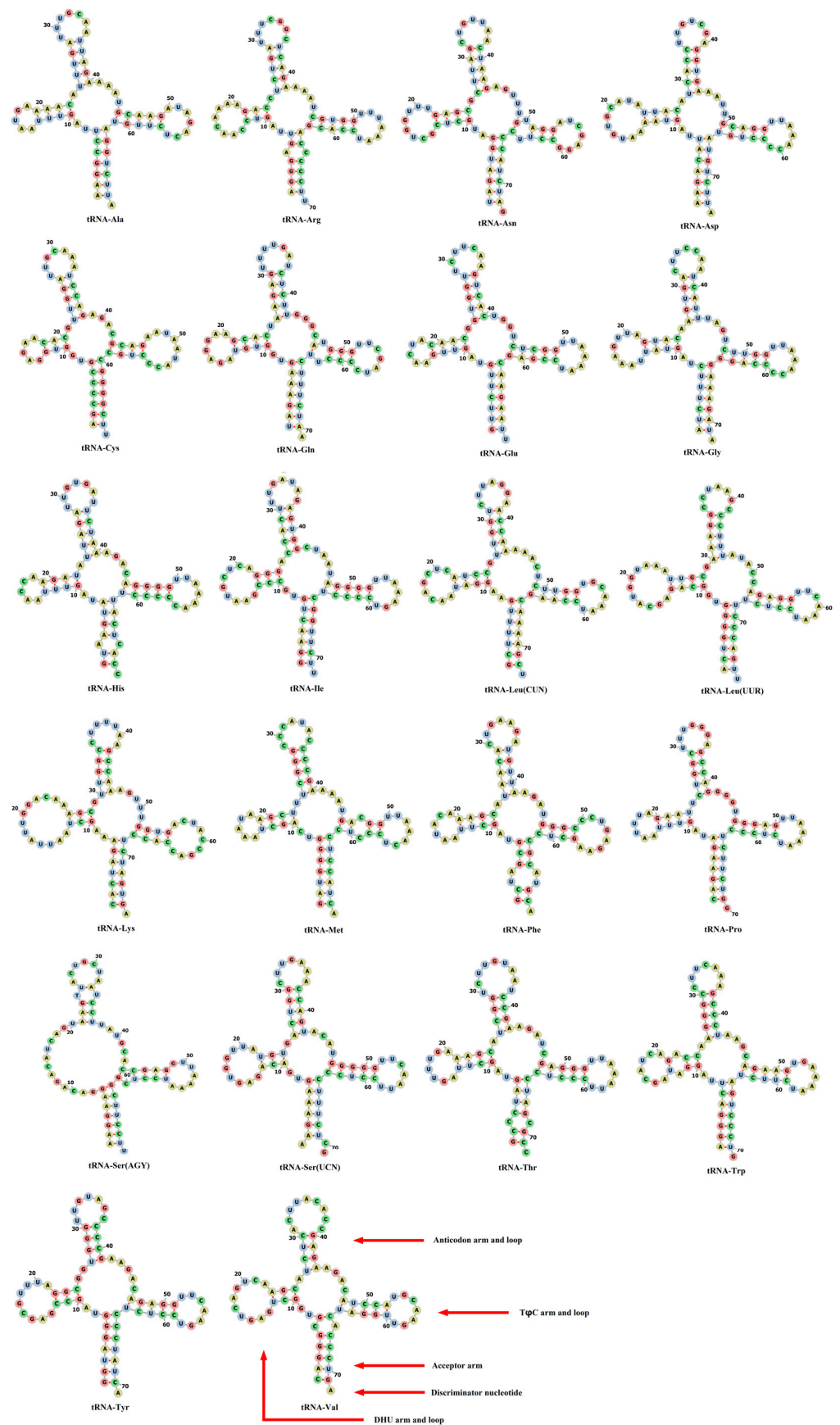


Figure 2. Secondary structures of mitochondrial tRNAs in *L. merzbacheri*.

3.4. Non-Coding Regions

The non-coding regions of mtDNA in *L. merzbacheri* included the O_L region and D-loop region. The O_L region resided between tRNA-Asn and tRNA-Cys with a whole length of 37 bp. The secondary structure of the O_L region was inferred using the Mfold tool (Figure 3). It consisted of 13 nt in the loop and 12 bp in the stem. The numbers of C and G in the stem were higher than those of A and T, while the reverse applied in the loop region, showing an obvious base asymmetry in this stem-loop structure. Similar to other vertebrates, the conserved sequence motif 3'-GGCCC-5' was present adjacent to the 3' end of the stem.

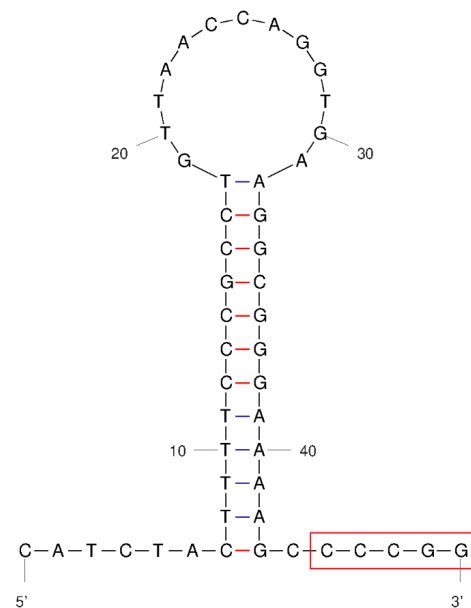
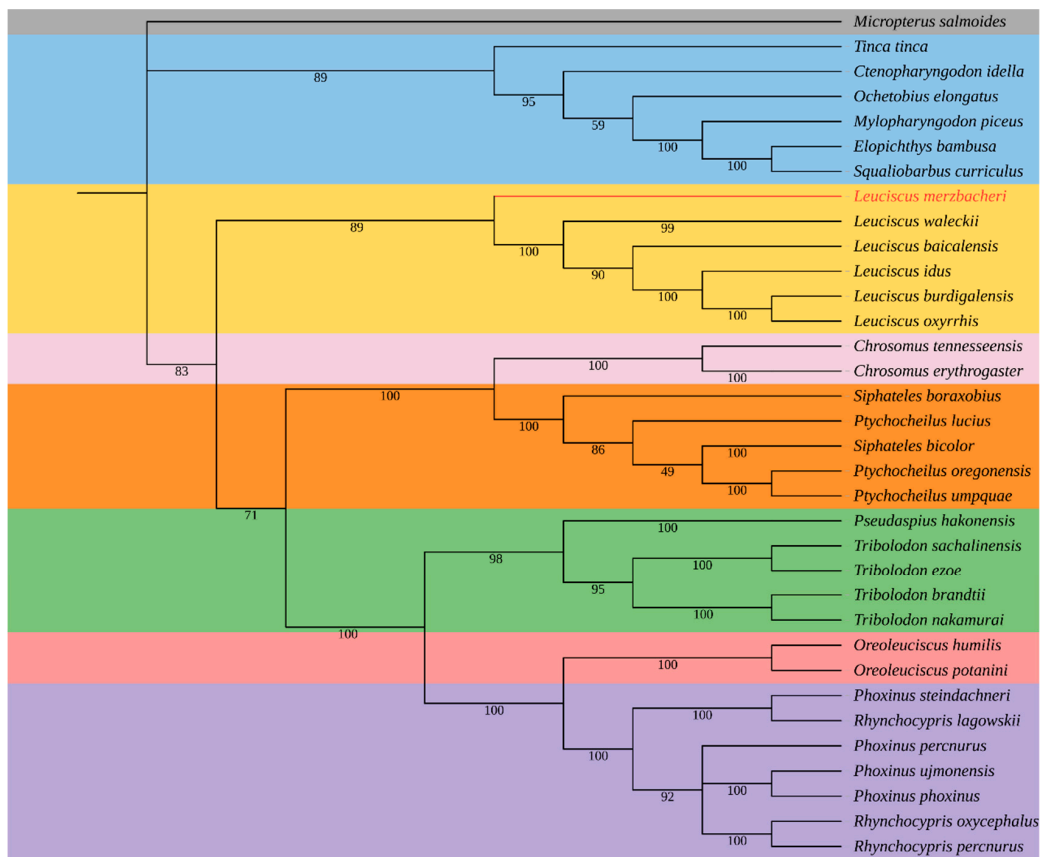


Figure 3. Stem-loop structure of O_L region in the mitogenome of *L. merzbacheri*. The conserved sequence region at the end of the stem is boxed.

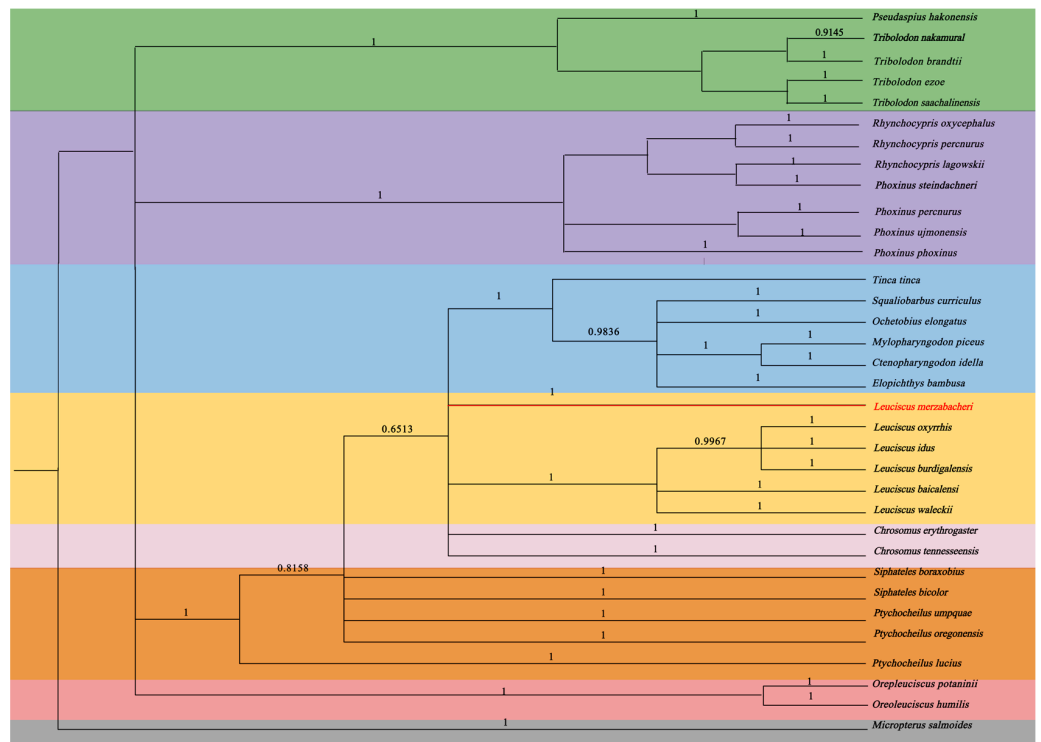
The length of entire mtDNA control region of *L. merzbacheri* was 929 bp. In this article, we recognized one extended termination associated sequence (ETAS), three enteral conserved elements (CSB-F, CSB-E and CSB-D), and three conserved elements (CSB-1, CSB-2 and CSB-3), respectively (Supplementary Figure S4). The ETAS was related to the synthesis termination of the D-loop, with a total length of 237 bp and the highest nucleotide variance rate. A termination-associated sequence 'TACAT' (53–57 bp) and its reverse complementary sequence 'ATGTA' (119–123 bp) were the key sequences of the ETAS domain, which made it form a stable hairpin structure. Meanwhile, the CD domain (333 bp) was the most conserved domain, and its typical central conserved block CSB-F was considered the distinguishing mark between ETAS and central CSB. CSB-1, CSB-2 and CSB-3 were found at the 3'-end of the control region with their iconic sequences 5'-CATCATTGAAAGACATA-3', 5'-CAAACCCCCTACCCC-3' and 5'-TGTCAAACCCCGAAACCAA-3', respectively. CSB-1 was at the start of the conserved sequence block and was relatively less conserved than CSB-2 and CSB-3.

3.5. Phylogenetic Relationships of *L. merzbacheri*

Two phylogenetic trees were constructed to analyze the evolutionary position of *L. merzbacheri* (Figure 4). The ML tree and BI tree exhibited similar topological structures, with species under the same taxonomic category roughly clustering into one clade. In the ML tree, *L. oxyrrhis* and *L. burdigalensis* first gathered together, and other *Leuciscus* species sequentially clustered outside. The BI tree was slightly different, with *L. oxyrrhis*, *L. idus* and *L. burdigalensis* clustering as a clade initially. A high bootstrap value and posterior probability were found in these two trees in which *L. merzbacheri* was located at the outermost level of the genus *Leuciscus* clade.



(A)



(B)

Figure 4. The ML (A) and BI (B) phylogenetic trees based on 13 concatenated PCG sequences. Numbers in the branches correspond to the bootstrap values (ML tree) and posterior probabilities (BI tree), respectively.

4. Discussion

The double-stranded circular mtDNA structure of *L. merzbacheri* was congruent with those of other vertebrates, and the total length was within the published length range of the Leuciscinae fishes' mitochondrial genomes (15,672–17,865 bp). A positive AT-skew value and a negative GC-skew value were detected, reflecting the AT richness of the mitochondrial genome as a whole [43]. Among 13 PCGs, seven of them ended with incomplete stop codons TA- or T-. The truncated stop codons were usual in the mitogenomes of teleosts [44] and might be modified to the TAA during the mitochondrial mRNA 3' polycistronic transcription and subsequent post-transcriptional polyadenylation [45].

The RSCU value means the number of occurrences of that codon in the genome. A value larger than 1.0 indicates that a given synonymous codon is preferred over the rest, less than 1.0 suggests a disfavored codon, and equivalent to 1.0 denotes no preference [46]. A total of 30 codons with RSCU values less than 1 were detected in all PCGs. Among them, the RSCU values of NNA codons were all greater than 1.0 only excluding UUA, which was positively correlated with the highest percentage of A at the third codon positions.

There were two copies of tRNA-Leu and tRNA-Ser in the mitogenome of *L. merzbacheri*, with the anti-codons UAA and UAG in the former, as well as the anti-codons UGA and GCU in the latter, respectively. It was speculated that the process of duplication, anticodon mutation and deletion of tRNA-Leu genes occurred during the evolution of the metazoan, and these two genes might have evolved independently for some time [47]. Only tRNA-Ser^{AGN} could not form the cloverleaf structure, just like many other bony fishes [48]. The phenomenon of a wobble base pair was quite common in tRNAs [49]. It was indicated that the base pair of swinging pairs accounts for approximately 7% of the tRNA stem regions in fish mitochondria [44]. Herein, the noncanonical base pair G-U had the highest frequency. It was deduced that G-U mispairing might be biologically relevant with nucleic acid secondary structures and the amino acid misincorporation [50,51] and could be repaired and corrected through RNA editing [52].

As indispensable components of ribosomes, rRNAs play an important role in life activities, and many studies have shown that the functions of RNA are closely related to their secondary structures [53]. In this study, the two-dimensional structure of 16S rRNA was much more complex than that of 12S rRNA, which suggested that the latter was more conserved than the former. Therefore, 12S rRNA has been regarded as a DNA meta-barcode for the interspecific identification of fishes [54]. Moreover, a conserved multibranching loop (CML) was found in domain VI of 16S rRNA, which was the core of the ribosomal peptidyl transferase center (PTC) in prokaryotic 23S rRNA, a homologue to mtDNA 16S rRNA [55]. It was speculated that it is functionally important in catalyzing chemical reactions of protein synthesis [44].

A recent study indicated that an East Asian group of the subfamily Leuciscinae evolved independently, distant from those in Europe, Siberia and North America [28]. *Leuciscus* is a monophyletic group [56], and the divergence time of *L. merzbacheri* was estimated to be earlier than other daces in China [19,57]. It was confirmed that *L. idus* and *L. baicalensis* were more closely related to each other, and *L. merzbacheri* was the most ancient species among three *Leuciscus* fishes in Xinjiang [16,17]. Geographic isolation plays an important role in biological evolution and speciation [58]. The Junggar Basin is the second largest inland basin in China, which is surrounded by the Tianshan Mountains, the Altai Mountains and the Junggar Boundary Mountains, and the Gurbantungu Desert is located in its center. In addition, this region is dominated by a typical continental arid climate with an average annual temperature of 3–7 °C and yearly precipitation of 200 mm, respectively [59]. The natural geographical barrier, coupled with drought and a rainless environment allow a species to evolve in its own unique direction, which possibly led to a notable heredity difference between *L. merzbacheri* and congeneric species.

5. Conclusions

In the current study, a total of a 16,609 bp mitogenome sequence of *L. merzbacheri* was obtained based on a high-throughput sequencing technique, with the gene positions and features consistent with those of the published mitochondrial genomes of *Leuciscus* fishes. The overall nucleotide composition showed a clear AT preference, which is very common in the mitogenomes of most bony fishes. The phylogenetic analysis revealed that *Leuciscus* has a monophyletic origin. *L. merzbacheri* was well clustered with other *Leuciscus* fishes, and it was perhaps the earliest diverging species of this clade. The complete mitogenome of *L. merzbacheri* would be a useful resource for studies on the molecular taxonomy, species identification and genetic diversity of the genus *Leuciscus*.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes15101284/s1>, Table S1: Information on the complete mitogenome sequences used in this study; Table S2: The nucleotide compositions and skewness of the *L. merzbacheri* mitogenome; Figure S1: Map of the mitochondrial genome of *L. merzbacheri*; Figure S2: The amino acid composition in 13 PCGs of the *L. merzbacheri* mitogenome; Figure S3: Secondary structures of rRNAs in the mitochondrial DNA of *L. merzbacheri*. (a) 12S rRNA. (b) 16S rRNA. Black bases indicate that they are the same as the template DNA, green bases indicate they are different from the template DNA, red bases indicate the inserted bases and blue bases indicate the reinsertion bases; Figure S4: Structural characteristics of the control region in *L. merzbacheri*.

Author Contributions: Conceptualization and writing—review and editing, T.Y.; investigation, methodology and writing—original draft preparation, Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Innovation and Entrepreneurship Training Program for College Students (No. 202310340056) and the Science and Technology Innovation Project of College Students in Zhejiang Province (No. 2023R411005).

Institutional Review Board Statement: The animal experiments were carried out in accordance with the guideline of the Ethics Committee for Animal Experimentation of Zhejiang Ocean University, Zhoushan, China, with approval code ZJOU-ECAE20220916 and approval date 16 September 2022.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in the GenBank of NCBI (<https://www.ncbi.nlm.nih.gov>; accessed on 7 January 2024) under the accession number OR992052.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Nelson, J.C.; Grande, T.C.; Wilson, M.V.H. *Fishes of the World*, 5th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2016.
- Chen, Y.Y. *Fauna Sinica, Osteichthyes, Cypriniformes II*; Science Press: Beijing, China, 1998.
- Guo, Y.; Zhang, R.M.; Cai, L.G. *Fishes of Xinjiang*; Xinjiang Science and Technology Press: Urumqi, China, 2012.
- Tu, Z.C. Study on Genetic Diversity of Endangered Fish: *Leuciscus merzbacheri* (Zugmayer 1912) and Introgression among the Three *Leuciscus* subspecies in Sailimu Lake of Xinjiang, China Using ISSR Markers. Master's Thesis, Shihezi University, Shihezi, China, 2011.
- Yue, P.Q.; Chen, Y.U. *China Red Data Book of Endangered Animals (Pisces)*; Science Press: Beijing, China, 1998.
- Cao, L.; Zhao, Y. *Leuciscus merzbacheri*. The IUCN Red List of Threatened Species. 2023. Available online: <https://www.iucnredlist.org/species/212785578/212785580> (accessed on 7 January 2024).
- Chen, X.Y. Studies on the skeleton of Leuciscine fishes of China, with particular reference to its significance in taxonomy. *Acta Zootax. Sin.* **1987**, *12*, 311–322.
- Guo, Y.; Cai, L.G.; Zhang, R.M.; Tu, E.X.; Zhang, B.P. Preliminary study on the biological properties of *Leuciscus merzbacheri* Zugmayer in the Sayram Lake, Xinjiang. *Arid Zone Res.* **2005**, *22*, 197–200.
- Huo, T.B.; Ma, B.; Tang, F.J.; Jiang, Z.F.; Cai, L.G.; Adakbek, K.J.; Liu, L.Z. The growth models of three kind of leuciscus from Xinjiang. *China J. Fish.* **2008**, *21*, 8–14.
- Wang, Y.N.; Li, S.Z.; Liu, C.L. Multivariate analysis on the morphological among three species of *Leuciscus* in Xinjiang. *J. Fish Sci. China* **2019**, *26*, 636–645.
- Li, S.Z. *Studies on Zoogeographical Divisions for Fresh Water Fishes of China*; Science Press: Beijing, China, 1981.

12. Song, M.B.; He, J.T.; Xie, C.G.; Cai, L.G.; Wang, S.; Liu, Y.; Shi, C.M. Evaluate status of marked release of *Leuciscus merzbacheri*. *Tianjin Agr. Sci.* **2021**, *27*, 41–45.
13. Song, M.B.; Hu, J.W.; Li, X.D.; Wang, X.; He, J.T.; Shi, C.M.; Wang, S.; Niu, J.G. Experimental starvation on *Leuciscus merzbacheri* larvae and determination of the point of no return. *J. Fish Res.* **2021**, *43*, 472–479.
14. Wang, J.J.; Hu, W.G.; Kong, L. Karyotype and banding pattern of *Leuciscus merzbacheri*. *China J. Zool.* **2010**, *45*, 120–126.
15. Lu, L.P.; Dong, Y.Y.; Hu, W.G.; Wang, J.J.; Chen, D.W.; Wang, X. The specific expression of isozymes in *Leuciscus merzbacheri* tissues. *China J. Zool.* **2012**, *47*, 105–111.
16. Hu, W.G.; Duan, Z.Y.; Wang, J.F.; Sheng, J.L.; Ma, R.L. Divergence and systematical evolution of three *Leuciscus* species in Xinjiang based on mitochondrial DNA control region sequences. *Acta Genet. Sin.* **2004**, *31*, 970–975.
17. Hu, W.G.; Duan, Z.Y.; Wang, J.F.; Sheng, J.L. Divergence and systematical evolution of three *Leuciscus* species in Xinjiang: Mitochondrial DNA cytochrome b sequences. *China J. Zool.* **2005**, *40*, 6–11.
18. Hu, W.G.; Ma, S.W.; Chen, C.F.; Wang, Y.Z.; Ren, Y.; Cao, X.D.; Sheng, J.L. Characterization of β -actin promoter of *Leuciscus merzbacheri*. *J. Anim. Vet. Adv.* **2012**, *11*, 3849–3854.
19. Meng, W.; Li, L.; Yuan, X.Q.; Zhou, Y.D. The complete mitochondrial genome of *Leuciscus merzbacheri* (Cypriniformes: Cyprinidae). *Mitochondrial DNA B* **2023**, *8*, 414–417. [[CrossRef](#)] [[PubMed](#)]
20. Klingenberg, M. The ADP, ATP shuttle of the mitochondrion. *Trends Biochem. Sci.* **1979**, *4*, 249–252. [[CrossRef](#)]
21. Taanman, J.W. The mitochondrial genome: Structure, transcription, translation and replication. *Biochim. Biophys. Acta-Bioenerg.* **1999**, *1410*, 103–123. [[CrossRef](#)] [[PubMed](#)]
22. Falkenberg, M.; Larsson, N.G.; Gustafsson, C.M. DNA replication and transcription in mammalian mitochondria. *Ann. Rev. Biochem.* **2007**, *76*, 679–699. [[CrossRef](#)]
23. Shadel, G.S.; Clayton, D.A. Mitochondrial DNA maintenance in vertebrates. *Ann. Rev. Biochem.* **1997**, *66*, 409–435. [[CrossRef](#)]
24. Avise, J.C.; Arnold, J.; Ball, R.M.; Bermingham, E.; Lamb, T.; Neigel, J.E.; Reeb, C.A.; Saunders, N.C. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* **1987**, *18*, 489–522. [[CrossRef](#)]
25. Wilson, M.R.; Stoneking, M.; Holland, M.M.; DiZinno, J.A.; Budowle, B. Guidelines for the use of mitochondrial DNA sequencing in forensic science. *Crime Lab Digest.* **1993**, *20*, 68–77.
26. Unajak, S.; Meesawat, P.; Anyamaneeratch, K.; Anuwareepong, D.; Srikulnath, K.; Choowongkamon, K. Identification of species (meat and blood samples) using nested-PCR analysis of mitochondrial DNA. *Afr. J. Biotechnol.* **2011**, *10*, 5670–5676.
27. Hai, S.; Cai, L.G.; Meng, W.; Yang, T.Y.; Ma, X.F. Divergence time estimation of fishes in Leuciscinae in Xinjiang. *China J. Fish.* **2014**, *27*, 12–16.
28. Wang, X.Z.; Liu, H.Z.; He, S.P.; Chen, Y.Y. Sequence analysis of cytochrome b gene indicates that East Asian group of cyprinid subfamily Leuciscinae (Teleostei: Cyprinidae) evolved independently. *Prog. Nat. Sci.* **2004**, *14*, 132–137. [[CrossRef](#)]
29. Imoto, J.M.; Saitoh, K.; Sasaki, T.; Yonezawa, T.; Adachi, J.; Kartavtsev, Y.P.; Miya, M.; Nishida, M.; Hanzawa, N. Phylogeny and biogeography of highly diverged freshwater fish species (Leuciscinae, Cyprinidae, Teleostei) inferred from mitochondrial genome analysis. *Gene* **2013**, *514*, 112–124. [[CrossRef](#)] [[PubMed](#)]
30. Pereira, C.S.A.; Aboim, M.A.; Ráb, P.; Collares-Pereira, M.J. Introgressive hybridization as a promoter of genome reshuffling in natural homoploid fish hybrids (Cyprinidae, Leuciscinae). *Heredity* **2014**, *112*, 343–350. [[CrossRef](#)] [[PubMed](#)]
31. Sambrook, J.; Russell, D. *Molecular Cloning—A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2000.
32. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2001**, *17*, 10–12. [[CrossRef](#)]
33. Hahn, C.; Bachmann, L.; Chevreur, B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—A baiting and iterative mapping approach. *Nucleic Acids Res.* **2013**, *41*, e129. [[CrossRef](#)]
34. Lohse, M.; Drechsel, O.; Kahlau, S.; Bock, R. OrganellarGenomeDRAW—A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **2013**, *41*, 575–581. [[CrossRef](#)]
35. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
36. Tamura, K.; Stecher, G.; Peterson, D.; Filipiński, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)]
37. Perna, N.T.; Kocher, T.D. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* **1995**, *41*, 353–358. [[CrossRef](#)]
38. Kück, P.; Longo, G.C. FASconCAT-G: Extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front. Zool.* **2014**, *11*, 81. [[CrossRef](#)]
39. Posada, D.; Crandall, K.A. Modeltest: Testing the model of DNA substitution. *Bioinformatics* **1998**, *14*, 817–818. [[CrossRef](#)] [[PubMed](#)]
40. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]
41. Ronquist, F.; Teslenko, M.; Mark, P.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; et al. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [[CrossRef](#)] [[PubMed](#)]

42. Hickson, R.E.; Simon, C.; Cooper, A.; Spicer, G.S.; Sullivan, J.; Penny, D. Conserved sequence motifs, alignment, and secondary structure for the third domain of animal 12S rRNA. *Mol. Biol. Evol.* **1996**, *13*, 150–169. [[CrossRef](#)] [[PubMed](#)]
43. Shioiri, C.; Takahata, N. Skew of mononucleotide frequencies, relative abundance of dinucleotides, and DNA strand asymmetry. *J. Mol. Evol.* **2001**, *53*, 364–376. [[CrossRef](#)] [[PubMed](#)]
44. Satoh, T.P.; Miya, M.; Mabuchi, K.; Nishida, M. Structure and variation of the mitochondrial genome of fishes. *BMC Genom.* **2016**, *17*, 719. [[CrossRef](#)]
45. Ojala, D.; Montoya, J.; Attardi, G. tRNA punctuation model of RNA procession in human mitochondria. *Nature* **1981**, *290*, 470–474. [[CrossRef](#)]
46. Behura, S.K.; Severson, D.W. Codon usage bias: Causative factors, quantification methods and genome-wide patterns: With emphasis on insect genomes. *Biol. Rev.* **2012**, *88*, 49–61. [[CrossRef](#)]
47. Higgs, P.G.; Jameson, D.; Jow, H.; Rattray, M. The evolution of tRNA-Leu genes in animal mitochondrial genomes. *J. Mol. Evol.* **2003**, *57*, 435–445. [[CrossRef](#)]
48. Miya, M.; Takeshima, H.; Endo, H.; Ishiguro, N.B.; Inoue, J.G.; Mukai, T.; Satoh, T.P.; Yamaguchi, M.; Kawaguchi, A.; Mabuchi, K.; et al. Major patterns of higher teleostean phylogenies: A new perspective based on 100 complete mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **2003**, *26*, 121–138. [[CrossRef](#)]
49. Rozov, A.; Demeshkina, N.; Khusainov, I.; Westhof, E.; Yusupov, M.; Yusupova, G. Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat. Commun.* **2016**, *7*, 10457. [[CrossRef](#)] [[PubMed](#)]
50. Müller, U.R.; Fitch, W.M. The biological significance of G-T/G-U mispairing in nucleic acid secondary structure. *J. Theor. Biol.* **1985**, *117*, 119–126. [[CrossRef](#)] [[PubMed](#)]
51. Zhang, Z.Q.; Shah, B.; Bondarenko, P.V. G/U and certain wobble position mismatches as possible main causes of amino acid misincorporations. *Biochemistry* **2013**, *52*, 8165–8176. [[CrossRef](#)] [[PubMed](#)]
52. Kumazawa, Y.; Nishida, M. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* **1993**, *37*, 380–398. [[CrossRef](#)] [[PubMed](#)]
53. Malhotra, A. *rRNA Structure*; John Wiley & Sons: Hoboken, NJ, USA, 2001.
54. Miya, M.; Sato, Y.; Fukunaga, T.; Sado, T.; Poulsen, J.Y.; Sato, K.; Minamoto, T.; Yamamoto, S.; Yamanaka, H.; Araki, H.; et al. MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *R. Soc. Open Sci.* **2015**, *2*, 150088. [[CrossRef](#)]
55. Polacek, N.; Mankin, A.S. The ribosomal peptidyl transferase center: Structure, function, evolution, inhibition. *Crit. Rev. Biochem. Mol. Biol.* **2005**, *40*, 285–311. [[CrossRef](#)]
56. Perea, S.; Böhme, M.; Zupančič, P.; Freyhof, J.; Šanda, R.; Özuluğ, M.; Abdoli, A.; Doadrio, I. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evol. Biol.* **2010**, *10*, 265. [[CrossRef](#)]
57. Pourshabanian, A.; Yazdani-Moghaddam, F.; Ghassemzadeh, F.; Mousavi-Sabet, H.; Rossi, G.; Aliabadian, M. Morpho-species of the genus *Leuciscus* Cuvier, 1816 (Teleostei: Leuciscinae) from Iran revisited using molecular approaches. *Zool. Middle East.* **2021**, *67*, 126–132. [[CrossRef](#)]
58. Bradburd, G.; Ralph, P.; Coop, G. Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution* **2013**, *67*, 3258–3273. [[CrossRef](#)]
59. Liu, J.Q.; Li, Z.C.; Qin, X.G. *Natural Environmental Evolution, Climate Change and Impact of Human Activities in Xinjiang Region*; China Water & Power Press: Beijing, China, 2014.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.